

P. Bui

RECEIVED Page 1 of 7

SEP 06 2000

TECH CENTER 1600/2900

1638

RAW SEQUENCE LISTING

DATE: 08/28/2000

PATENT APPLICATION: US/09/360,934A

TIME: 14:47:45

Input Set : A:\Chir0158.app

Output Set: N:\CRF3\08282000\I360934A.raw

ENTERED

```
3 <110> APPLICANT: Covacci, Antonello
4   Bugnoli, Massimo
5   Telford, John
6   Macchia, Giovanni
7   Rappuoli, Rino
9 <120> TITLE OF INVENTION: Helicobacter Pylori Cytotoxin Proteins Useful For
10  Vaccines And Diagnostics
12 <130> FILE REFERENCE: CHIR0158
14 <140> CURRENT APPLICATION NUMBER: 09/360,934A
15 <141> CURRENT FILING DATE: 1999-07-26
17 <150> PRIOR APPLICATION NUMBER: 08/466,662
18 <151> PRIOR FILING DATE: 1995-06-06
20 <150> PRIOR APPLICATION NUMBER: 08/256,858
21 <151> PRIOR FILING DATE: 1994-10-21
23 <160> NUMBER OF SEQ ID NOS: 10
25 <170> SOFTWARE: PatentIn Ver. 2.1
27 <210> SEQ ID NO: 1
28 <211> LENGTH: 27
29 <212> TYPE: DNA
30 <213> ORGANISM: Artificial Sequence
32 <220> FEATURE:
33 <223> OTHER INFORMATION: Description of Artificial Sequence: Novel
34   Sequence
36 <400> SEQUENCE: 1
37 gcaagcttat cgatgtcgac tcgagct
40 <210> SEQ ID NO: 2
41 <211> LENGTH: 3960
42 <212> TYPE: DNA
43 <213> ORGANISM: Helicobacter pylori
45 <400> SEQUENCE: 2
46 aaaaagaaag gaagaaaatg gaaatacaac aaacacaccg caaatcaat cgccctctgg 60
47 ttctctctgc tttagtagga gcattagtca gcatcacacc gcaacaaagt catgccgcct 120
48 ttttcacaac cgtgatcatt ccagccattg ttgggggtat cgctacaggc accgctgtag 180
49 gaacggtctc agggcttctt agctgggggc tcaaacaagc cgaagaagcc aataaaaccc 240
50 cagataaacc cgataaagtt tggcgcattc aagcaggaaa aggctttaat gaattcccta 300
51 acaaggaata cgacttatac agatcccttt tatccagtaa gattgatgga ggttgggatt 360
52 gggggaatgc cgctagggcat tattgggtca aaggcgggca acagaataag cttgaagtgg 420
53 atatgaaaga cgctgtaggg acttatacct tatcagggct tagaaaacttt actgggtggg 480
54 atttagatgt caatatgcaa aaagccactt tacgcttggg ccaattcaat ggcaattctt 540
55 ttacaagcta taaggatagt gctgatcgca ccacgagagt ggatttcaac gctaaaaata 600
56 tctcaattga taattttgta gaaatcaaca atcgtgtggg ttctggagcc gggaggaaag 660
57 ccagctctac ggttttgact ttgcaagctt cagaagggat cactagcgat aaaaacgctg 720
58 aaatttctct ttatgatggt gccacgctca atttgcttc aagcagcgtt aaattaatgg 780
59 gtaatgtgtg gatgggccgt ttgcaatacg tgggagcgta ttggccctt tcatacagca 840
60 cgataaacac ttcaaaagta acaggggaag tgaattttaa ccacctcact gttggcgata 900
61 aaaacgccgc tcaagcgggc attatcgcta ataaaaagac taatttggc acactggatt 960
62 tgtggcaaa gcccgggtta aacattatcg ctctccaga aggtggctat aaggataaac 1020
```

RAW SEQUENCE LISTING
 PATENT APPLICATION: US/09/360,934A
 DATE: 08/28/2000
 TIME: 14:47:45

Input Set : A:\Chir0158.app
 Output Set: N:\CRF3\08282000\I360934A.raw

```

63 ccaataatac cctttctcaa agtgggtgcta aaaacgacaa aaatgaaagc gctaaaaaacg 1080
64 acaaaacaaga gagcagtcac aataatagta acactcaggt cattaaccca cccaatagtg 1140
65 cgcaaaaaaac agaagttcaa cccacgcaag tcattgatgg gccttttgcg ggcggcaaaag 1200
66 acacggttgt caatatcaac cgcacatcaac ctaacgctga tggcacgatt agagtgggag 1260
67 ggttttaaagc ttctcttacc accaatgcgg ctcatttgca tatcgcaaaa ggcgggtgtca 1320
68 atctgtccaa tcaagcgagc gggcgctctc ttatagtggg aaatctaact gggaaatatca 1380
69 ccggtgatgg gcctttaaga gtgaataatc aagtgggtgg ctatgctttg gcaggatcaa 1440
70 gcgcgaattt tgagttaaag gctggtacgg ataccacaaa cggcacagcc acttttaata 1500
71 acgatattag tctgggaaga tttgtgaatt taaaggtgga tgctcatata gctaatttta 1560
72 aaggtattgg tacgggtaac ggtggtttca acaccttaga ttttagtggc gttacagaca 1620
73 aagtcaatat caacaagctc attacggctt ccactaatgt ggcggttaaa aacttcaaca 1680
74 ttaatgaatt gattgttaaa accaatggga taagtgtggg ggaatatact cattttagcg 1740
75 aagatatagg cagtcaatcg cgcacatcaa ccgtagcgtt ggaaactggc actaggtcac 1800
76 ttttctctgg ggtgtgtaaa tttaaaggtg gcgaaaaatt ggttatagat gagttttact 1860
77 atagcccttg gaattatttt gacgctagaa atattaaaaa tgttgaaatc accaataaac 1920
78 ttgcttttgg acctcaagga agtccttggg gcacatcaaa acttatgttc aataatctaa 1980
79 ccctaggtca aaatgcggtc atggattata gccaattttc aaatttaacc attcaagggg 2040
80 atttcatcaa caatcaaggc actatcaact atctggtccg aggtgggaaa gtggcaacct 2100
81 taagcgtagg caatgcagca gctatgatgt ttaataatga tatagacagc gcgaccgat 2160
82 ttacaaacc gctcatcaag attaacagcg ctcaagatct cattaaaaat acagaacatg 2220
83 ttttattgaa agcgaaaaatc attggttatg gtaatgttcc tacaggtacc aatggcatta 2280
84 gtaatgttaa tctagaagag caattcaaag agcgcttagc cttttataac aacaataacc 2340
85 gcatggatac ttgtgtgggt cgaaatactg atgacattaa agcatgcggt atggctatcg 2400
86 gcgatcaaaag gatgggtgaac aacctgaca attacaagta tcttatcggt aaggcatgga 2460
87 aaaatatagg catcagcaaa acagctaatg gctctaaaaa ttcggtgtat tattaggca 2520
88 attctacgcc tactgagaat ggtggcaata ccacaaattt acccacaacc accactagca 2580
89 atgcacgttc tgccaacaac gcccttgcac aaaacgctcc ttctgctcaa cctagtgcta 2640
90 ctccataatt agtcgctatc aatcagcatg attttggcac tattgaaagc gtgtttgaat 2700
91 tggctaaccg ctctaaagat attgacacgc tttatgctaa ctacggcgct caaggcaggg 2760
92 atctcttaca aaccttattg attgatagcc atgatgcggg ttatgccaga aaaatgattg 2820
93 atgctacaag cgctaataag atcaccaagc aattgaatac ggccactacc actttaaaac 2880
94 acatagccag tttagagcat aaaaccagcg gcttacaaac tttgagcttg agtaatgcga 2940
95 tgattttaaa ttctcgttta gtcaatctct ccaggagaca caccacccat attgactcgt 3000
96 tcgccaacag ctacaaagct ttaaaagacc aaaaattcgc ttcttttaga agcgcgagcag 3060
97 aagtgttgta tcaatttgcc cctaaatatg aaaaacctac caatgtttgg gctaacgcta 3120
98 ttgggggaac gagcttgaat aatggctcta acgcttcatt gtatggcaca agcgcgggcg 3180
99 tagacgctta ccttaacggg caagtggag ccattgtggg cggttttgga agctatggtt 3240
100 atagctcttt taataatcgt gcgaactccc ttaactctgg ggccaataac actaattttg 3300
101 gcgtgtatag cgtatatttt gccaacagc atgaatttga ctttgaagct caaggggcac 3360
102 tagggagcga tcaatcaagc ttgaatttca aaagcgctct attacaagat ttgaatcaaa 3420
103 gctatcatta cttagcctat agcgctgcaa caagagcgag ctatggttat gacttcgct 3480
104 tttttaggaa cgcttttagt ttaaaaccaa gcgtgggtgt gagctataac catttaggtt 3540
105 caaccaactt taaaagcaac agcaccaatc aagtggcttt gaaaaatggc tctagcagtc 3600
106 agcattttatt caacgctagc gctaattgtg aagcgcgcta ttattatggg gacacttcat 3660
107 actctacat gaatgctgga gttttacaag agttcgctca tgttggctct aataacgccg 3720
108 cgtctttaaa cacttttaaa gtgaatgcg ctcgcaaccc tttaaatacc catgccagag 3780
109 tgatgatggg tggggaatta aaattagcta aagaagtgtt tttgaatttg ggcgttggtt 3840
110 atttgacaa tttgatttcc aatataggcc atttcgcttc caatttagga atgaggtata 3900
111 gtttctaaat accgctctta aaccatgct caaagcatgg gtttgaaatc ttacaaaaca 3960

```

RAW SEQUENCE LISTING DATE: 08/28/2000
 PATENT APPLICATION: US/09/360,934A TIME: 14:47:45

Input Set : A:\Chir0158.app
 Output Set: N:\CRF3\08282000\I360934A.raw

```

114 <210> SEQ ID NO: 3
115 <211> LENGTH: 1296
116 <212> TYPE: PRT
117 <213> ORGANISM: Helicobacter pylori
119 <400> SEQUENCE: 3
120 Met Glu Ile Gln Gln Thr His Arg Lys Ile Asn Arg Pro Leu Val Ser
121 1 5 10 15
123 Leu Ala Leu Val Gly Ala Leu Val Ser Ile Thr Pro Gln Gln Ser His
124 20 25 30
126 Ala Ala Phe Phe Thr Thr Val Ile Ile Pro Ala Ile Val Gly Gly Ile
127 35 40 45
129 Ala Thr Gly Thr Ala Val Gly Thr Val Ser Gly Leu Leu Ser Trp Gly
130 50 55 60
132 Leu Lys Gln Ala Glu Glu Ala Asn Lys Thr Pro Asp Lys Pro Asp Lys
133 65 70 75 80
135 Val Trp Arg Ile Gln Ala Gly Lys Gly Phe Asn Glu Phe Pro Asn Lys
136 85 90 95
138 Glu Tyr Asp Leu Tyr Arg Ser Leu Leu Ser Ser Lys Ile Asp Gly Gly
139 100 105 110
141 Trp Asp Trp Gly Asn Ala Ala Arg His Tyr Trp Val Lys Gly Gly Gln
142 115 120 125
144 Gln Asn Lys Leu Glu Val Asp Met Lys Asp Ala Val Gly Thr Tyr Thr
145 130 135 140
147 Leu Ser Gly Leu Arg Asn Phe Thr Gly Gly Asp Leu Asp Val Asn Met
148 145 150 155 160
150 Gln Lys Ala Thr Leu Arg Leu Gly Gln Phe Asn Gly Asn Ser Phe Thr
151 165 170 175
153 Ser Tyr Lys Asp Ser Ala Asp Arg Thr Thr Arg Val Asp Phe Asn Ala
154 180 185 190
156 Lys Asn Ile Ser Ile Asp Asn Phe Val Glu Ile Asn Asn Arg Val Gly
157 195 200 205
159 Ser Gly Ala Gly Arg Lys Ala Ser Ser Thr Val Leu Thr Leu Gln Ala
160 210 215 220
162 Ser Glu Gly Ile Thr Ser Asp Lys Asn Ala Glu Ile Ser Leu Tyr Asp
163 225 230 235 240
165 Gly Ala Thr Leu Asn Leu Ala Ser Ser Ser Val Lys Leu Met Gly Asn
166 245 250 255
168 Val Trp Met Gly Arg Leu Gln Tyr Val Gly Ala Tyr Leu Ala Pro Ser
169 260 265 270
171 Tyr Ser Thr Ile Asn Thr Ser Lys Val Thr Gly Glu Val Asn Phe Asn
172 275 280 285
174 His Leu Thr Val Gly Asp Lys Asn Ala Ala Gln Ala Gly Ile Ile Ala
175 290 295 300
177 Asn Lys Lys Thr Asn Ile Gly Thr Leu Asp Leu Trp Gln Ser Ala Gly
178 305 310 315 320
180 Leu Asn Ile Ile Ala Pro Pro Glu Gly Gly Tyr Lys Asp Lys Pro Asn
181 325 330 335
183 Asn Thr Pro Ser Gln Ser Gly Ala Lys Asn Asp Lys Asn Glu Ser Ala
184 340 345 350

```

RAW SEQUENCE LISTING DATE: 08/28/2000
 PATENT APPLICATION: US/09/360,934A TIME: 14:47:45

Input Set : A:\Chir0158.app
 Output Set: N:\CRF3\08282000\I360934A.raw

```

186 Lys Asn Asp Lys Gln Glu Ser Ser Gln Asn Asn Ser Asn Thr Gln Val
187           355           360           365
189 Ile Asn Pro Pro Asn Ser Ala Gln Lys Thr Glu Val Gln Pro Thr Gln
190           370           375           380
192 Val Ile Asp Gly Pro Phe Ala Gly Gly Lys Asp Thr Val Val Asn Ile
193 385           390           395           400
195 Asn Arg Ile Asn Thr Asn Ala Asp Gly Thr Ile Arg Val Gly Gly Phe
196           405           410           415
198 Lys Ala Ser Leu Thr Thr Asn Ala Ala His Leu His Ile Gly Lys Gly
199           420           425           430
201 Gly Val Asn Leu Ser Asn Gln Ala Ser Gly Arg Ser Leu Ile Val Glu
202           435           440           445
204 Asn Leu Thr Gly Asn Ile Thr Val Asp Gly Pro Leu Arg Val Asn Asn
205           450           455           460
207 Gln Val Gly Gly Tyr Ala Leu Ala Gly Ser Ser Ala Asn Phe Glu Phe
208 465           470           475           480
210 Lys Ala Gly Thr Asp Thr Lys Asn Gly Thr Ala Thr Phe Asn Asn Asp
211           485           490           495
213 Ile Ser Leu Gly Arg Phe Val Asn Leu Lys Val Asp Ala His Thr Ala
214           500           505           510
216 Asn Phe Lys Gly Ile Asp Thr Gly Asn Gly Gly Phe Asn Thr Leu Asp
217           515           520           525
219 Phe Ser Gly Val Thr Asp Lys Val Asn Ile Asn Lys Leu Ile Thr Ala
220           530           535           540
222 Ser Thr Asn Val Ala Val Lys Asn Phe Asn Ile Asn Glu Leu Ile Val
223 545           550           555           560
225 Lys Thr Asn Gly Ile Ser Val Gly Glu Tyr Thr His Phe Ser Glu Asp
226           565           570           575
228 Ile Gly Ser Gln Ser Arg Ile Asn Thr Val Arg Leu Glu Thr Gly Thr
229           580           585           590
231 Arg Ser Leu Phe Ser Gly Gly Val Lys Phe Lys Gly Gly Glu Lys Leu
232           595           600           605
234 Val Ile Asp Glu Phe Tyr Tyr Ser Pro Trp Asn Tyr Phe Asp Ala Arg
235           610           615           620
237 Asn Ile Lys Asn Val Glu Ile Thr Asn Lys Leu Ala Phe Gly Pro Gln
238 625           630           635           640
240 Gly Ser Pro Trp Gly Thr Ser Lys Leu Met Phe Asn Asn Leu Thr Leu
241           645           650           655
243 Gly Gln Asn Ala Val Met Asp Tyr Ser Gln Phe Ser Asn Leu Thr Ile
244           660           665           670
246 Gln Gly Asp Phe Ile Asn Asn Gln Gly Thr Ile Asn Tyr Leu Val Arg
247           675           680           685
249 Gly Gly Lys Val Ala Thr Leu Ser Val Gly Asn Ala Ala Met Met
250           690           695           700
252 Phe Asn Asn Asp Ile Asp Ser Ala Thr Gly Phe Tyr Lys Pro Leu Ile
253 705           710           715           720
255 Lys Ile Asn Ser Ala Gln Asp Leu Ile Lys Asn Thr Glu His Val Leu
256           725           730           735
258 Leu Lys Ala Lys Ile Ile Gly Tyr Gly Asn Val Ser Thr Gly Thr Asn

```

RAW SEQUENCE LISTING DATE: 08/28/2000
 PATENT APPLICATION: US/09/360,934A TIME: 14:47:45

Input Set : A:\Chir0158.app
 Output Set: N:\CRF3\08282000\I360934A.raw

259		740		745		750
261	Gly	Ile	Ser	Asn	Val	Asn
262		755		760		765
264	Leu	Tyr	Asn	Asn	Asn	Arg
265		770		775		780
267	Asp	Asp	Ile	Lys	Ala	Cys
268	785			790		795
270	Asn	Asn	Pro	Asp	Asn	Tyr
271		805		810		815
273	Ile	Gly	Ile	Ser	Lys	Thr
274		820		825		830
276	Leu	Gly	Asn	Ser	Thr	Pro
277		835		840		845
279	Pro	Thr	Asn	Thr	Thr	Ser
280		850		855		860
282	Gln	Asn	Ala	Pro	Phe	Ala
283	865			870		875
285	Ile	Asn	Gln	His	Asp	Phe
286		885		890		895
288	Asn	Arg	Ser	Lys	Asp	Ile
289		900		905		910
291	Gly	Arg	Asp	Leu	Gln	Thr
292		915		920		925
294	Tyr	Ala	Arg	Lys	Met	Ile
295		930		935		940
297	Gln	Leu	Asn	Thr	Ala	Thr
298	945			950		955
300	His	Lys	Thr	Ser	Gly	Leu
301		965		970		975
303	Leu	Asn	Ser	Arg	Leu	Val
304		980		985		990
306	Asp	Ser	Phe	Ala	Lys	Arg
307		995		1000		1005
309	Ser	Leu	Glu	Ser	Ala	Ala
310		1010		1015		1020
312	Glu	Lys	Pro	Thr	Asn	Val
313	1025			1030		1035
315	Asn	Asn	Gly	Ser	Asn	Ala
316		1045		1050		1055
318	Ala	Tyr	Leu	Asn	Gly	Gln
319		1060		1065		1070
321	Tyr	Gly	Tyr	Ser	Ser	Phe
322		1075		1080		1085
324	Ala	Asn	Asn	Thr	Asn	Phe
325		1090		1095		1100
327	His	Glu	Phe	Asp	Phe	Glu
328	1105			1110		1115
330	Ser	Leu	Asn	Phe	Lys	Ser
331		1125		1130		1135

VERIFICATION SUMMARY DATE: 08/28/2000
PATENT APPLICATION: US/09/360,934A TIME: 14:47:46

Input Set : A:\Chir0158.app
Output Set: N:\CRF3\08282000\I360934A.raw